

ACT WORKING PAPER 2016 (04)

Using Incoming Student Information to Identify Students At-Risk of Not Returning to Their Initial Institution in Year Two

Justine Radunzel, PhD

October, 2016

ACT Working Paper Series

ACT working papers document preliminary research. The papers are intended to promote discussion and feedback before formal publication. The research does not necessarily reflect the views of ACT.



Justine Radunzel is a principal research scientist in Statistical and Applied Research specializing in postsecondary outcomes research and validity evidence for the ACT® test.

Abstract

As pushes for increased accountability in higher education continue, postsecondary institutions are interested in identifying early on students who are at risk of leaving their institution. With this in mind, this study sought to identify incoming first-year student information (such as that available on the ACT student record) that postsecondary institutions might use for determining students who are at-risk of leaving their institution in year two. Specifically, student characteristics were examined in relation to two types of attrition for the institution – students dropping out of college and students transferring to another institution – in comparison to students returning in year two.

Data were available for more than 630,000 ACT-tested 2014 high school graduates who enrolled in college in fall 2014 at nearly 1,150 two- and four-year institutions. Initial and subsequent enrollment was tracked using National Student Clearinghouse data. Data on student-level characteristics evaluated included academic preparation and achievement measures; college intentions about living on campus, enrolling full-time, and working while in college; educational goals; the number of college preferences met by the initial institution; the distance between home and initial institution attended; and demographic characteristics. Students' college intentions, college preference matches, and distance from home were included in the study as possible proxies for barriers to social and academic integration at the initial institution attended. Hierarchical multinomial regression models accounting for institution attended were used to estimate retention and attrition rates. Results were examined by type of institution.

Study findings suggest that multiple academic and non-academic factors are useful for predicting student attrition. First, at both two- and four-year institutions, students who were less academically prepared for college were more likely than those who were better prepared to drop

out of college. Academic readiness was also negatively related to transfer at four-year institutions but was somewhat positively related to transfer at two-year institutions. College intentions also played a role in identifying who was likely to leave their initial institution. For example, students who indicated that they planned to work more hours while in college were more likely to drop out of college than those intending to work fewer hours. Additionally, the fewer the number of college preferences met by the initial institution attended the more likely a student was to drop out or transfer to another institution. Attending an institution farther away from home was also associated with higher transfer rates. Unfortunately, even after statistically controlling for academic measures and other student characteristics, students from certain underserved demographic groups (e.g., first-generation students and economically disadvantaged students) continued to be somewhat more likely than their peers to drop out. In secondary analyses, retention and attrition rates were found to vary across college majors and by institutional characteristics.

Study findings illustrate how institutions can use incoming student information from the ACT record to help identify students who are at-risk of leaving their institution, allowing for the opportunity to intervene early with these students. The ACT student record contains additional data elements beyond those examined in this study that can help institutions build multidimensional models of student success in order to better identify students who might benefit from additional academic and student support services upon entering college.

Using Incoming Student Information to Identify Students At-Risk of Not Returning to Their Initial Institution in Year Two

Over the past decade, postsecondary institutions have been under considerable pressure to increase their retention and degree completion rates while maintaining equal opportunity and diversity in student enrollments. Recent statistics on a national sample of students from the 2008 college freshman cohort suggest that only 60% of students who initially enroll in four-year institutions complete a degree within 150% of normal time from their initial institution attended (Kena et al., 2016). The corresponding percentage for students initially enrolling in two-year institutions is considerably lower at 28%. Other research suggests that the largest share of students leave their initial institution during their first two years (Bradburn, 2002; Tinto, 2012). In response to pushes for increased accountability in higher education, postsecondary institutions continue to invest resources to better understand the academic and non-academic factors associated with student retention with the ultimate goal of improving degree completion rates on their campus. Specifically, institutions are interested in identifying early on students who are at risk of dropping out of college or transferring to another institution so that they can implement interventions and provide resources that address and support the needs of these students.

According to Tinto (1975; 1993), student' chances of being retained at an institution are influenced by students' pre-entry attributes, academic goals and commitments, institutional experiences, academic and social integration into the college environment, and external commitments. One pre-entry characteristic that has been found to be positively related to student retention is academic readiness, with academic readiness often being measured by standardized test scores, high school grade point average (HSGPA), and taking higher-level coursework in high school (e.g., ACT, 2013; Kopp & Shaw, 2016; Lotkowski, Robbins, & Noeth, 2004; Mattern & Patterson, 2009). In terms of high school coursework taken, a study by Adelman

(2006) found that the highest level of high school mathematics coursework is an important factor associated with bachelor's degree completion. A policy brief by Achieve (2008) suggests that the reason high school mathematics preparation is so important for college success is related to the higher-order thinking and critical reasoning skills that students learn beginning in Algebra I and continue to build upon in subsequent higher-level mathematics courses. Students who develop these skills are better equipped for their future career pathways, whichever path they may choose to follow.

Student demographic characteristics have also been found to be related to student retention. White students have generally been found to have higher retention rates than underrepresented racial/ethnic minority students (e.g., Mattern & Patterson, 2009). However, there have been some studies that have found that once academic readiness measures and other student characteristics are statistically controlled for in the models that racial/ethnic gaps in retention rates are substantially reduced (ACT, 2010; Radunzel & Noble, 2012) or even reversed (Ishitani, 2016; Kopp & Shaw, 2016). As for gender differences in retention rates, study findings have been mixed. However, in national studies, female students are generally more likely than male students to persist and complete a college degree (Kena et al., 2016).

Other demographic characteristics that have been found to be related to student retention are socioeconomic status and parent's education level. Lower socioeconomic levels have been found to be associated with lower retention rates, even after controlling for academic readiness measures and other student characteristics (ACT, 2010; Kopp & Shaw, 2016; Radunzel & Noble, 2012). Lower-income students are generally more likely than their peers to have non-academic obligations, such as the need to work and/or have family responsibilities that can influence their study habits and chances of persisting in college (Engle & Tinto, 2008). First-generation students

(those whose parents have no college experience) are also less likely than their peers to persist and complete a degree (Ishitani, 2016; Kopp & Shaw, 2016). The gap in retention rates by parental education is likely due in part to first-generation students generally being less likely than their peers to be academically prepared for college, to have early exposure to and knowledge about the college environment, and to have the guidance and support at home that can help contribute to student success in college (Saenz, Hurtado, Barrera, Wolf, & Yeung, 2007).

The extent to which students are committed to attaining a college degree is another factor positively related to students' chances of being retained (Allen, Robbins, Casillas, & Oh, 2008), as are various measures of academic and social integration (Ishitani, 2016). Some choices that can help foster academic and social integration include: living on campus, participating in campus activities, enrolling full-time, and attending an institution that matches students' preferences (Bowman & Denson, 2014; Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006). In contrast, having to work many hours off campus while going to college can prevent students from "fully engaging in the college environment" (Kuh et al., 2006).

Attending a college farther from home has also been found to be negatively related to social integration and college adjustment as well as being positively related to homesickness (e.g., Brooks & DuBois, 1995; Fisher, Murray, & Frazer, 1985; Tognoli, 2003). Moreover, a recent study by Mattern, Wyatt, and Shaw (2013) found that students who attended an institution farther away from home had greater chances of leaving their initial institution and transferring to another institution than students who attended an institution closer to home, even after statistically controlling for academic readiness and demographic characteristics.

Current Study

Building on prior research, the primary objective of the current multi-institutional study was to identify incoming student information that might be useful for determining early on students who are at-risk of leaving their initial institution in year two while also differentiating between two types of student attrition that may occur from the institution's perspective: drop out (i.e., students who are not enrolled at any institution) and transfer (i.e., students who enroll at another institution). Student characteristics evaluated included academic preparation and achievement measures; college intentions about living on campus, enrolling full-time, and working while in college; educational goals; the number of college preferences met by the initial institution; the distance between home and initial institution attended, and demographic characteristics. Students' college intentions, college preference matches, and distance from home were included in the study as possible proxies for barriers to social and academic integration at the initial institution attended. In an earlier ACT study (2014), many of these same student-level characteristics were found to be related to dropout and transfer in descriptive analyses. The current study will extend beyond these descriptive findings to develop a model of students' chances of being retained at year two in relation to multiple student characteristics simultaneously. Findings from this study will help to illustrate how postsecondary institutions – both four- and two-year institutions – might use readily available student information from the ACT record to build and augment their multidimensional models of student success that help to identify early on students on their campuses that are likely to leave their institutions.

Secondary objectives for the study included examining the variation in retention and attrition rates across college majors and by institutional characteristics. The former is considered a secondary objective because college majors were available for only a subset of students in the

current sample. Prior research has found that retention rates and degree completion rates vary across academic major, even after statistically controlling for academic readiness, financial aid, and demographic characteristics (Wohlgemuth, Whalen, Sullivan, Nading, Shelley, & Wang, 2007).¹ We evaluated the college major/retention results in the current study to illustrate that institutions might want to evaluate these relationships on their own campuses. Having this additional information could lead to greater insights on ways their retention strategies might be improved and become more personalized to keep students on their campuses and help them to succeed.

Beyond statistically controlling for student characteristics, the current study also examined retention and attrition rates by institutional characteristics to further explain the variability in rates observed across institutions. Prior studies have found retention rates to vary by institution type, size, and selectivity (Bowen, Chingos, & McPherson, 2009; Kopp & Shaw, 2016). Moreover, institutional characteristics can also play a role in social integration (Tinto, 1993).

Data

Sample

Initial data consisted of roughly 1,275,485 students who graduated from high school in 2014, took the ACT, and enrolled in a postsecondary institution in fall 2014 (76% attended a four-year institution and 24% attended a two-year institution; ACT, 2015a). Initial enrollment in fall 2014 and subsequent enrollment data in fall 2015 were obtained from the National Student Clearinghouse (NSC).² The study sample was restricted to institutions that had at least 50% of

¹ At the institution being studied, students' selected their initial college major in one of seven undergraduate colleges. The seven colleges included: Agriculture; Business; Design; Education; Engineering; Family and Consumer Science; and the College of Liberal Arts and Sciences.

² Data from NSC accounts for 95% of all enrollments in Title IV, degree-granting institutions in the nation.

their incoming students who had taken the ACT to ensure that the ACT-tested population reasonably represented the institution's incoming first-time entering student population,³ restricting the sample to 920,508 students.

Analyses were done separately by institution type, where type was determined at the time of initial enrollment. Eighty-three percent of the students in the restricted sample enrolled in a four-year institution in fall 2014 (766,503 students from 877 four-year institutions; 154,005 students from 272 two-year institutions). Even though ACT test scores are generally not required for admissions to two-year institutions, in states that administer the ACT statewide, most if not all public high school graduates will have ACT scores and other student characteristics from the ACT record available for use by institutions. Therefore, we included in this study results for ACT-tested students from both two- and four-year institutions.

Data for students' demographic characteristics, high school coursework taken, grades earned in those courses, educational goals, college intentions and preferences, and official ACT test scores were obtained from the ACT student record. The self-reported information was provided by students at the time they registered to take the ACT. If students took the ACT more than once, only data from the most recent ACT administration was used. The final sample used for this study was comprised of 527,090 students (or 69%) from the four-year sample and 106,219 students (or 69%) from the two-year sample who completed all questionnaire items analyzed in this study.⁴

³ Institutional enrollment counts of first-time, degree-seeking undergraduate students for fall 2014 were obtained from the Integrated Postsecondary Education Data System (IPEDS).

⁴ The variables with the highest percentage of missing responses included parental education level (14%), number of hours plan to work while in college (13% to 14%), and the number of college preferences met by the initial institution attended (18% to 21%).

Table 1 provides a description of the institutions included in the four-year and two-year samples. Seventy percent of the institutions were from the East and Midwest regions for both samples. Compared to two-year institutions, a higher percentage of four-year institutions were private institutions.

Table 1. *Description of Institutions in Samples*

College characteristics	Four-year sample (N=877 institutions)		Two-year sample (N=272 institutions)	
	<i>n</i>	Pct.	<i>n</i>	Pct.
Control				
Private	528	60	4	1
Public	349	40	268	99
HBCU				
Yes	47	5	4	1
No	830	95	268	99
Size				
Under 1,000	73	8	12	4
1,000 – 4,999	452	52	168	62
5,000 – 9,999	132	15	56	21
10,000 – 19,999	113	13	28	10
20,000 and above	107	12	8	3
Region				
East	310	35	96	35
Midwest	311	35	96	35
Southwest	111	13	37	14
West	145	17	43	16
Admissions selectivity				
Highly selective	88	10	0	0
Selective	239	27	0	0
Traditional	408	47	0	0
Liberal	77	9	0	0
Open	65	7	272	100

Note. Characteristics for the postsecondary institutions were obtained from IPEDS, except for admissions selectivity. Admission selectivity was self-reported by institutions on the ACT Institutional Data Questionnaire as defined by the typical high school class ranks of their accepted freshmen: the majority of freshmen at highly selective schools are in the top 10%, selective in the top 25%, traditional in the top 50%, liberal in the top 75% of their high school class (ACT, 2015b). Institutions with open admissions policies accept all high school graduates to the limit of capacity. The average number of ACT-tested students per institution was 601 students for the four-year sample and 389 students for the two-year sample.

Measures

Study Outcome

The study outcome was whether a student returned during the fall of year two (fall 2015) to the same initial institution attended in year one (fall 2014). This variable was coded into the following three distinct categories to allow for the examination of two types of attrition from an institution's perspective:

- returned to initial institution
- transferred to another institution
- dropped out (not enrolled in college).

These are point-in-time definitions of “transfer” and “dropout”; it is possible that students classified as such will reenroll at some point in the future.

Predictors

Variables examined as predictors of student retention in the current study are listed below. Demographic characteristics included:

- gender
- race/ethnicity (categorized as African American; American Indian; Hispanic; Asian, Pacific Islander; Multiracial; and White)
- highest parental education level (categorized as no college experience; some college experience or earned an associate's degree; earned a bachelor's degree; and earned a master's, doctorate, or professional degree (e.g., M.D., J.D.))

- median household income associated with student’s residential zip code (categorized as \$43,315 or less; \$43,316 — \$61,580, more than \$61,580). The median household income by zip code was based on 2006 to 2010 data from the American Community Survey.⁵

Academic preparation and achievement measures included:

- ACT Composite score (the rounded arithmetic average of the four subject area scores in English, Mathematics, Reading, and Science). ACT Composite score was evaluated as a continuous variable (1–36) as well as a categorical variable (1 to 15; 16 to 19; 20 to 23; 24 to 27; 28 to 36).
- HSGPA (students’ self-reports of their coursework taken in up to 23 specific courses in English, mathematics, social studies, and science, and the grades earned in those courses). HSGPA was evaluated as a continuous variable (0.00–4.00) as well as a categorical variable (0.00 to 1.99; 2.00 to 2.49; 2.50 to 2.99; 3.00 to 3.49; 3.50 to 3.74; 3.75 to 4.00).
- highest mathematics course taken in high school (categorized as Calculus; Trigonometry or other advanced math beyond Algebra II; Algebra II; and below Algebra II).⁶

Prior studies have shown that students report high school coursework and grades accurately relative to information provided in their official high school transcripts (Sanchez & Buddin, 2016; Shaw & Mattern, 2009).

⁵ Data for median household income by zip code was obtained from the following site: <http://www.psc.isr.umich.edu/dis/census/Features/tract2zip/> provided by the Michigan Population Studies Center. The zip code of the high school attended was used in cases where a student’s residential zip code was missing. Note that the median household income was used instead of student’s self-reported annual family income due to a high percentage of students not reporting this information (22% to 25%). In comparison, median household income per zip code could be determined for most students (99%). Among students with values for both self-reported annual family income and median household income per zip code, the Spearman correlation coefficient between these two variables was 0.39 for the four-year sample and 0.29 for the two-year sample.

⁶ Students were asked to indicate courses that they had taken, were currently taking, or planned to take before graduating from high school. In this study, courses that students had taken, were currently taking, and planned to take were coded as taken in high school.

Some predictors that were included in this study to serve as proxies for possible barriers to social and academic integration included students' college intentions, number of college preferences met by initial institution attended, and distance between a student's home and college attended. College intentions and educational goals included plans for:

- living on campus (categorized as yes; no)
- enrolling as a full-time student (categorized as yes; no)
- the number of hours plan to work per week while in college (categorized as 0; 1 – 10; 11 – 20; 21 – 30; or 31 or more hours)
- the highest level of education expected to complete (categorized as associate's degree or below; bachelor's degree; beyond a bachelor's degree; or other).⁷

Students were asked about their college preferences on the following characteristics:

- type of institution⁸
- state location of the institution
- size of the institution.⁹

Students' college preferences were then compared to the characteristics of the initial institution attended. The number of college preferences met by the initial institution was the predictor considered in the retention models. The values for this variable ranged from 0 to 3 matches.

The distance between a student's home address and college address was another predictor included in the models. Distance was calculated using a SAS function that returns the geodetic

⁷ Associate's degree or below included the following: a business/technical or certificate program or an associate's degree. Beyond a bachelor's degree included the following: a master's degree, a doctoral degree, or a professional level degree (e.g., M.D., J.D.).

⁸ Students' preferences on institution type included the following possibilities: four-year public, four-year private, or two-year institution.

⁹ Students' preferences on institution size were categorized using similar groupings as those used in IPEDS data. These categories included: less than 1,000 students; 1,000 to 5,000 students; 5,000 to 10,000 students; 10,000 to 20,000 students; or 20,000 or more students.

distance in miles between two zip code locations.¹⁰ To account for (1) the heavily right-skewed distribution of distance values and (2) cases where students attended a college that had the same zip code as their home address (i.e., distance equaled 0), the distance variable was transformed using the log base 10 transformation as follows: LOG base 10 (distance + 1).¹¹ The transformed distance variable was classified into the following categories: 0.00 to 1.49; 1.50 to 2.24, and 2.25 or more. These categories translate to the following approximate categories based on miles from home: 0 to 30 miles; 31 miles to 174 miles; and 175 or more miles. The transformed distance variable was also examined as a continuous predictor.

The data used to address the secondary study objectives of examining retention and attrition rates by college major and institutional characteristics came from NSC, IPEDS, and ACT's Institutional Data Questionnaire (IDQ). Six-digit Classification of Instruction Program (CIP) major codes for each term of enrollment were available in the NSC data for only a subset of the sample (445,115 students or 84% for the four-year sample and 84,674 students or 80% for the two-year sample). In the analyses, majors were classified according to the first two digits of the CIP codes. Thirty-six different major clusters were represented (see Table A-4 for the list). Retention and attrition rates were reported only for major clusters that included at least 100 students. The institutional characteristics that are presented in Table 1 were those that were considered as predictors in these secondary analyses (except region). Because all or nearly all of the two-year institutions were public institutions and had open admissions policies, institution size was the only institution-level variable evaluated in the model for the two-year sample.

¹⁰ Distance was calculated based on a student's residential zip code obtained from their ACT record and the postsecondary institution's zip code obtained from IPEDS using the ZIPCITYDISTANCE function in SAS. The centroid of each zip code is used in the distance calculations.

¹¹ Distance from home is 0 miles, 9 miles, 99 miles, and 999 miles when the transformed distance variable is 0, 1, 2, and 3, respectively.

Method

Due to the nested structure of the data (i.e., students clustered within institutions), hierarchical regression models were developed to predict retention from the student and institutional characteristics. A hierarchical multinomial regression model was used for the three-category retention outcome.¹² Returning to the initial institution was the base category for the study outcome. Hierarchical models provide two general types of estimates: (1) fixed effects, which estimate the value of the parameter at a typical institution, and (2) variance estimates, which describe the variability of the parameter estimates across institutions. In these models, intercepts were allowed to vary randomly across institutions. For both the four- and two-year samples, single-predictor models, as well as a multiple-predictor model based on all predictors jointly, were developed.

For each variable, the odds ratio (OR) was reported as a means to compare the strength of the predictor-outcome relationships among student characteristics. Two ORs of attrition compared to the base category were estimated: the OR of dropping out vs. returning to the initial institution and the OR of transferring to another institution vs. returning to the initial institution. The OR represents the odds of experiencing the outcome (e.g., dropping out compared to returning) for a certain subgroup of students (e.g., female students, students taking Calculus in high school), compared to the odds of experiencing the outcome for another subgroup of students (e.g., male students, students not taking Calculus in high school; the latter group is often referred to as the referent group).¹³

¹² THE GLIMMIX procedure for generalized mixed models, available in SAS 9.2, with the Laplace estimation method and generalized logit link was used to fit the models.

¹³ For a multinomial outcome, the odds of experiencing a particular outcome (e.g., dropping out) is the ratio of the probability of experiencing the outcome (e.g., dropping out) to the probability of experiencing the base outcome (e.g., returning to the initial institution).

In comparison to members in the referent group, an OR greater than 1.0 indicates that students in the subgroup of interest are generally more likely to experience the outcome of interest, whereas an OR less than 1.0 indicates that they are less likely to do so. An OR estimated from a single-predictor model is labeled as an *unadjusted OR*. An OR estimated from a multiple-predictor model is labeled as an *adjusted OR*, because the OR reflects the effect of taking into account other student characteristics. The 99% confidence interval for the OR provides an indication of whether the relationship is statistically significant at the 0.01 level (that being when the interval does not include the null value of 1.0). In addition to ORs, retention and attrition rates by student and institutional characteristics were reported to help provide context for the practical significance of the findings, especially in light of the relatively large sample size. From the multiple-predictor models, retention and attrition rates by student and institutional characteristics were estimated using the fixed effect parameter estimates from the hierarchical models and holding all other predictors in the model constant at the sample means.

Results

Description of Study Samples

Table 2 contains descriptive statistics on student demographics for the study samples. Female students made up more than 50% of each sample (57% for the four-year sample and 55% for the two-year sample). Nearly two-thirds of the students in each sample were White students. As compared to the four-year sample, the two-year sample was comprised of a higher percentage of students from less affluent neighborhoods (40% vs. 24%) and of students whose parents had no college experience (26% vs. 13%).

Table 2. *Description of Student Demographics by Study Samples*

Student characteristic	Four-year sample (N=527,090)		Two-year sample (N=106,219)	
	<i>n</i>	Pct.	<i>n</i>	Pct.
Gender				
Male	225,858	43	47,575	45
Female	301,232	57	58,644	55
Race/ethnicity				
African American	63,416	12	17,119	16
American Indian	3,104	1	898	1
Hispanic	55,760	11	11,651	11
Asian	23,156	4	2,302	2
Pacific Islander	1,003	<1	206	<1
Multiracial	20,708	4	3,905	4
White	359,943	68	70,138	66
Median household income*				
< \$43,316	124,003	24	42,382	40
\$43,316 to \$61,580	179,800	34	37,091	35
> \$61,580	223,287	42	26,746	25
Highest parental education level				
No college	68,063	13	27,391	26
Some college	133,977	25	39,418	37
Bachelor's degree	174,978	33	26,683	25
Graduate degree	150,072	28	12,727	12

* Median household income is based students' residential zip code.

Table 3 provides descriptive statistics on students' academic preparation and achievement measures. Students in the four-year sample tended to have higher ACT Composite scores and HSGPAs than did those in the two-year sample. Students in the four-year sample were nearly two times more likely to take a Calculus course in high school than were those in the two-year sample.

Table 3. *Description of Academic Preparation and Achievement Measures by Study Samples*

Student characteristic	Four-year sample (N=527,090)		Two-year sample (N=106,219)	
	<i>n</i>	Pct.	<i>n</i>	Pct.
ACT Composite score				
1 to 15	21,857	4	21,252	20
16 to 19	93,812	18	40,046	38
20 to 23	161,225	31	30,704	29
24 to 27	145,323	28	11,764	11
28 to 36	104,873	20	2,453	2
HSGPA				
0.00 to 1.99	3,361	1	5,406	5
2.00 to 2.49	18,413	3	15,270	14
2.50 to 2.99	59,408	11	24,788	23
3.00 to 3.49	151,414	29	33,046	31
3.50 to 3.74	110,094	21	14,023	13
3.75 to 4.00	184,400	35	13,686	13
Highest mathematics course				
Calculus	258,306	49	27,071	25
Trig/Other Adv. math	229,216	43	52,498	49
Algebra II	37,459	7	23,687	22
Below Algebra II	2,109	<1	2,963	3

Note. The mean ACT Composite scores is 23.3 for the four-year sample and 19.0 for the two-year sample. The mean HSGPA is 3.46 for the four-year sample and 3.02 for the two-year sample.

Table 4 provides descriptive statistics on students' college intentions and educational goals. Students in the four-year sample were more likely than students in the two-year sample to indicate that they planned to live on campus (81% vs. 46%). For both samples, 90% or more of students indicated that they planned to enroll in college as a full-time student. A higher percentage of students from the two-year sample indicated that they planned to work more than 10 hours per week while attending college than did those in the four-year sample (67% vs. 49%). Students from the four-year sample were nearly two times more likely than those from the two-year sample to indicate that they aspired to earn a post-baccalaureate degree (51% vs. 28%).

Table 4. *Description of College Intentions and Plans by Study Samples*

College intentions/plans	Four-year sample (N=527,090)		Two-year sample (N=106,219)	
	<i>n</i>	Pct.	<i>n</i>	Pct.
Live on campus				
Yes	427,197	81	48,786	46
No	99,893	19	57,433	54
Enroll full-time				
Yes	517,608	98	95,153	90
No	9,482	2	11,066	10
Hours plan to work				
None	113,931	22	10,140	10
1-10	153,395	29	24,515	23
11-20	196,664	37	46,048	43
21-30	53,625	10	20,700	19
31 or more	9,475	2	4,816	5
Educational goals				
Beyond bachelor's degree	266,348	51	29,251	28
Bachelor's degree	250,537	48	64,196	60
Other	3,372	1	1,912	2
Associate's or below	6,833	1	10,860	10

Table 5 provides descriptive statistics on whether students' college preferences on type of institution, state location, and size of institution were met by their initial institution attended. For both samples, a relatively high percentage (80% or more) of students' college preferences on state location were matched by those of their initial institution. In comparison, only one-third or fewer students attended an institution that matched their preference on the size of the student body. There was a large difference between the two samples in the percentage of students that had their preferred institution type matched by their initial institution: 78% for the four-year sample versus only 18% for the two-year sample.¹⁴ The four-year sample was nearly two times

¹⁴ For the two-year sample, more than 80% of students indicated that they preferred to attend a four-year institution. For the four-year sample, less than 2% of students indicated that they preferred to attend a two-year institution. Students' preferences on institution type included the following possibilities: four-year public, four-year private, or two-year institution.

more likely than the two-year sample to have two or more of their college preferences met by their initial institution attended (73% vs. 37%).

Table 5. *Description of College Preferences Met by Study Samples*

College preference met by initial institution attended	Four-year sample (N=527,090)		Two-year sample (N=106,219)	
	<i>n</i>	Pct.	<i>n</i>	Pct.
Type of institution				
Met	409,082	78	19,633	18
Not met	118,008	22	86,586	82
State location				
Met	423,539	80	91,390	86
Not met	103,551	20	14,829	14
Institution size				
Met	171,826	33	29,861	28
Not met	355,264	67	76,358	72
Number met				
0	29,786	6	9,904	9
1	115,049	22	57,758	54
2	257,367	49	32,545	31
3	124,888	24	6,012	6

Table 6 provides descriptive statistics on college enrollment characteristics. Three-fourths of students in the four-year sample and more than 90% of students in the two-year sample initially enrolled in an in-state institution. Students in the two-year sample were nearly three times more likely than those in the four-year sample to enroll in an institution closer to home (76% vs. 27%; within 30 miles from home).

Table 6. *Description of College Enrollment Characteristics by Study Samples*

College characteristics	Four-year sample (N=527,090)		Two-year sample (N=106,219)	
	<i>n</i>	Pct.	<i>n</i>	Pct.
In-state				
Yes	396,362	75	100,233	94
No	130,728	25	5,986	6
LOG(Distance + 1) ¹				
0.00 to 1.49	143,356	27	80,717	76
1.50 to 2.24	242,931	46	20,284	19
2.25 and higher	140,803	27	5,218	5

¹ The mean of the transformed distance from home variable was 1.78 for the four-year sample and 1.18 for the two-year sample. The categories for the transformed distance variable translate to the following approximate categories based on miles from home: 0 to 30 miles; 31 miles to 174 miles; and 175 or more miles.

The typical retention rate was 76% for the four-year and 60% for the two-year sample, after accounting for variability across institutions (Figure 1).¹⁵ These rates are consistent with those recently reported nationally for first-time, full-time, degree-seeking students from the 2013 cohort (80% at four-year institutions and 61% at two-year institutions; Kena et al., 2016). The typical dropout rate was lower for the four-year sample than for the two-year sample (13% vs. 30%), whereas the typical transfer rate was similar between the two samples (11% vs. 10%). For the four-year sample, among those who transferred in year two, 57% transferred to another four-year institution and 43% transferred to a two-year institution. For the two-year sample, the corresponding percentages were 62% transferred to a four-year institution and 38% transferred to another two-year institution.

¹⁵ These rates were based on the null random-intercept model that did not include any student or institution characteristics. In comparison, the observed retention rate was 77% for the four-year sample and 60% for the two-year sample. The modeled dropout and transfer rates were also comparable to the corresponding observed rates for each sample.

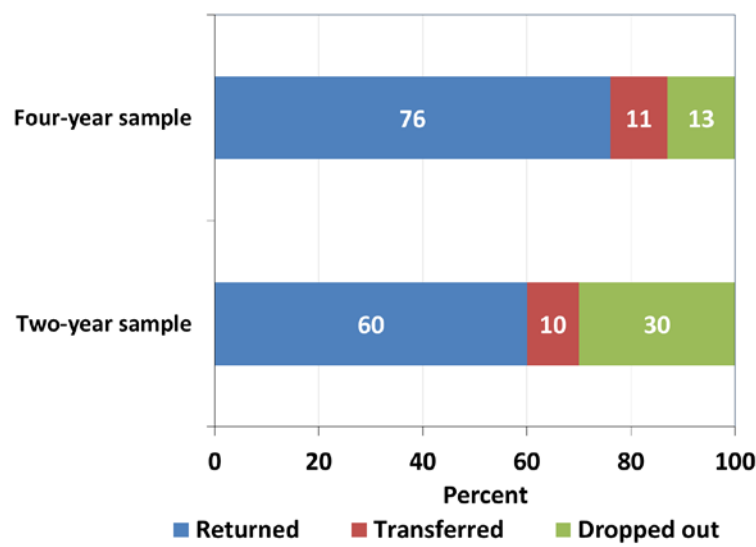


Figure 1. Modeled retention and attrition rates by sample, accounting for institution attended

Retention and attrition rates varied by student characteristics. Because all of the student characteristics remained statistically significant at the 0.01 level in the multiple-predictor models, we bypass discussing the results from the single-predictor models, and instead we discuss in detail in the next section the findings from the multivariate results.¹⁶

Multivariate Results by Student Characteristics

For both samples, the multiple-predictor multinomial model included demographic characteristics, academic preparation and achievement measures, college intentions and plans, number of college preferences met by initial institution attended, and distance from home. Modeled retention and attrition rates and adjusted ORs for the dropped out vs. returned and the transferred vs. returned comparisons are provided in Table 7 for the four-year sample and in Table 8 for the two-year sample. Corresponding results by student demographic characteristics are provided in Table A-1 for the four-year sample and Table A-2 for the two-year sample in the Appendix.

¹⁶ The overall p values for each predictor in the multiple-predictor models were < 0.001 for the four-year sample and < 0.0001 for the two-year sample.

Table 7. *Multivariate Results for First-to-Second Year Retention for the Four-Year Sample*

Student characteristics	Modeled rates			Dropped out vs. returned			Transferred vs. returned		
	Dropped out	Trans- ferred	Returned	OR	99% CI		OR	99% CI	
ACT Composite score									
1 to 15	14	12	74	1.44	1.35	1.54	1.97	1.82	2.13
16 to 19	13	13	75	1.32	1.26	1.38	1.97	1.87	2.08
20 to 23	12	11	77	1.17	1.12	1.22	1.76	1.68	1.85
24 to 27	11	10	79	1.07	1.02	1.11	1.43	1.36	1.49
28 to 36	11	7	82						
HSGPA									
0.00 to 1.99	21	12	67	3.03	2.72	3.38	1.84	1.59	2.13
2.00 to 2.49	20	13	67	2.94	2.78	3.11	2.01	1.87	2.15
2.50 to 2.99	17	13	70	2.32	2.22	2.41	1.83	1.75	1.91
3.00 to 3.49	14	11	75	1.74	1.68	1.80	1.55	1.49	1.60
3.50 to 3.74	11	10	79	1.30	1.25	1.34	1.29	1.24	1.34
3.75 to 4.00	9	8	83						
Highest math course									
Calculus	11	10	79	0.74	0.64	0.84	0.94	0.78	1.13
Trig/Other Adv Math	12	10	78	0.77	0.67	0.88	1.00	0.83	1.20
Alg II	13	10	76	0.92	0.80	1.05	1.03	0.85	1.24
Below Alg II	15	10	75						
Intend to live on campus									
Yes	11	10	78	0.93	0.91	0.96	1.05	1.04	1.12
No	12	10	78						
Intend to enroll full-time									
Yes	11	10	78	0.84	0.79	0.90	1.13	1.02	1.24
No	14	9	78						
Hours plan to work per week									
None	10	10	81	0.46	0.43	0.49	0.84	0.76	0.93
1-10	10	10	80	0.51	0.47	0.54	0.84	0.76	0.92
11-20	12	10	77	0.62	0.58	0.66	0.93	0.85	1.02
21-30	15	11	74	0.80	0.74	0.86	1.03	0.93	1.13
31 or more	18	10	71						
Educational plans									
Beyond bachelor's	11	10	78	0.99	0.91	1.07	1.36	1.21	1.52
Bachelor's degree	12	10	78	1.00	0.92	1.08	1.31	1.17	1.46
Other	12	9	79	1.07	0.93	1.24	1.17	0.97	1.42
Associate's or below	12	8	80						
Number of college preferences met									
0	12	12	76	1.18	1.12	1.25	1.30	1.23	1.38

Student characteristics	Modeled rates			Dropped out vs. returned			Transferred vs. returned		
	Dropped out	Transferred	Returned	OR	99% CI		OR	99% CI	
1	12	11	77	1.14	1.10	1.18	1.25	1.20	1.30
2	11	10	79	1.05	1.02	1.08	1.08	1.04	1.11
3	11	9	80						
LOG(Distance + 1) ¹									
0.00 to 1.49	12	8	80	1.08	1.05	1.12	0.63	0.60	0.65
1.50 to 2.24	11	11	78	1.02	0.99	1.05	0.93	0.90	0.96
2.25 and higher	11	12	77						

Note. Italics indicate referent group. Adjustment was made for all student characteristics included in the table, as well as for the demographic characteristics provided in Table A-1. OR = odds ratio; CI = confidence interval.

¹The categories for the transformed distance from home variable translate to the following approximate categories based on miles from home: 0 to 30 miles; 31 miles to 174 miles; and 175 or more miles.

The modeled retention and attrition rates were estimated holding all other variables constant at their sample means. For most of the predictors, because many of the variables were highly related to one another, the adjusted ORs from the multiple-predictors models were smaller than the unadjusted ORs from the single-predictor models (data not shown).

Variability estimates for the random intercepts from the null and multivariate models are provided in Table A-3. Based on McFadden's R^2 analog (McFadden, 1974), the percentage of variance explained by the multiple-predictor model that accounted for institution attended and the student-level predictors was 8% for the four-year sample and 7% for the two-year sample.¹⁷ These R^2 estimates are consistent with those reported in other studies on first-to-second year retention (D'Amico & Dika, 2013; Kopp & Shaw, 2016). It is important to note that pseudo R^2 values for binary or multinomial outcomes are typically smaller in magnitude than R^2 values for continuous outcomes.

¹⁷ The R^2 estimate attributed to the multiple student-level predictors in the model was 0.03 for the four-year sample and 0.06 for the two-year sample when comparing the log likelihoods between the multiple-predictor model and the intercept only model, conditional on the inclusion of the random intercepts.

Table 8. *Multivariate Results for First-to-Second Year Retention for the Two-Year Sample*

Student characteristics	Modeled rates			Dropped out vs. returned			Transferred vs. returned		
	Dropped out	Trans- ferred	Returned	OR	99% CI		OR	99% CI	
ACT Composite score									
1 to 15	33	7	59	1.24	1.07	1.45	0.74	0.62	0.90
16 to 19	29	8	63	1.01	0.87	1.17	0.78	0.66	0.93
20 to 23	26	9	64	0.91	0.79	1.05	0.88	0.74	1.04
24 to 27	26	9	64	0.90	0.78	1.05	0.87	0.73	1.04
28 to 36	28	10	62						
HSGPA									
0.00 to 1.99	45	8	47	4.09	3.68	4.55	1.25	1.05	1.49
2.00 to 2.49	40	8	53	3.22	2.95	3.50	1.15	1.02	1.30
2.50 to 2.99	33	8	59	2.44	2.25	2.63	1.06	0.95	1.18
3.00 to 3.49	28	9	64	1.85	1.71	1.99	1.06	0.96	1.16
3.50 to 3.74	22	9	69	1.38	1.26	1.50	1.07	0.96	1.18
3.75 to 4.00	17	9	73						
Highest math course									
Calculus	28	9	63	0.83	0.74	0.93	1.21	0.97	1.56
Trig/Other Adv Math	28	9	63	0.80	0.71	0.89	1.16	0.93	1.44
Alg II	31	8	62	0.91	0.82	1.02	1.03	0.83	1.29
Below Alg II	33	7	60						
Intend to live on campus									
Yes	29	11	61	1.06	1.01	1.10	1.56	1.46	1.65
No	29	7	64						
Intend to enroll full-time									
Yes	28	9	63	0.77	0.72	0.81	1.16	1.03	1.31
No	34	7	59						
Hours plan to work per week									
None	23	11	66	0.49	0.45	0.55	1.13	0.95	1.34
1-10	25	9	66	0.56	0.51	0.61	0.96	0.82	1.13
11-20	29	8	63	0.66	0.61	0.72	0.93	0.79	1.09
21-30	34	8	59	0.82	0.75	0.90	0.89	0.76	1.06
31 or more	38	8	54						
Educational plans									
Beyond bachelor's	29	10	62	0.94	0.88	1.01	1.43	1.25	1.63
Bachelor's degree	28	8	63	0.91	0.85	0.97	1.23	1.09	1.40
Other	32	7	61	1.06	0.91	1.21	1.09	0.83	1.45
Associate's or below	31	7	62						
Number of college preferences met									
0	34	10	56	1.57	1.42	1.74	1.74	1.46	2.08

Student characteristics	Modeled rates			Dropped out vs. returned			Transferred vs. returned		
	Dropped out	Transferred	Returned	OR	99% CI		OR	99% CI	
1	28	9	63	1.15	1.06	1.25	1.45	1.24	1.70
2	28	8	64	1.11	1.02	1.21	1.29	1.10	1.52
3	26	7	67						
LOG(Distance + 1) ¹									
0.00 to 1.49	29	7	64	1.03	0.93	1.14	0.33	0.30	0.38
1.50 to 2.24	28	14	58	1.11	1.00	1.23	0.69	0.61	0.78
2.25 and higher	25	19	56						

Note. Italics indicate referent group. Adjustment was made for all student characteristics included in the table, as well as for the demographic characteristics provided in Table A-2. OR = odds ratio; CI = confidence interval.

¹The categories for the transformed distance from home variable translate to the following approximate categories based on miles from home: 0 to 30 miles; 31 miles to 174 miles; and 175 or more miles.

For both samples, students who were better prepared academically were generally more likely than those who were less prepared to return to their initial institution in the fall of year 2. However, the primary source of the attrition (not enrolled vs. transferred) differed slightly between the two samples and depended on the academic preparation/achievement measure. For the four-year sample, students' chances of dropping out or transferring to another institution were greater for those with lower ACT Composite scores and HSGPAs as compared to those with higher achievement levels (e.g., adjusted OR = 1.1 to 1.4 for dropping out and adjusted OR = 1.4 to 2.0 for transferring based on ACT Composite score; Table 7). In comparison for the two-year sample, retention rates were lower among students with lower achievement levels primarily because these students were more likely to drop out (e.g., adjusted OR = 1.4 to 4.1 based on HSGPA; Table 8). Two-year students with higher ACT Composite scores were slightly more likely than those with lower scores to transfer to another institution (adjusted OR = 0.7 to 0.9 for lower vs. higher scoring students). As a result, modeled retention rates were somewhat more comparable between the lowest and highest ACT score groups for the two-year sample than for the four-year sample (59% to 62% vs. 74% to 82%, respectively). For both samples, students

who had taken higher-level mathematics coursework in high school (e.g., Calculus, Trigonometry, or another advanced math course) were more likely than those whose highest mathematics course was below Algebra II to return to their initial institution primarily because the former groups were less likely to drop out of college (adjusted OR = 0.7 to 0.8 for dropped out vs. returned).

Students' college intentions also played a role in identifying who was at risk of not returning to their initial institution. For both samples, students who indicated that they planned to work greater hours per week while attending college were less likely than those who planned to work fewer hours to return to the same institution (71% vs. 81% for four-year sample and 54% vs. 66% for two-year sample comparing more than 30 hours to 0 hours). The higher retention rates among those planning to work fewer hours was primarily attributed to these students being less likely to drop out of college (adjusted OR = 0.5 to 0.8 for both samples). From a practical significance perspective, this predictor had little to no effect on attrition rates due to transferring to another institutions (adjusted OR = 0.8 to 1.1 for both samples).

Students' intentions of living on campus and enrolling full-time had larger effects for the two-year sample than for the four-year sample. For the two-year sample, modeled retention rates were lower for students who indicated that they planned to live on campus than for those who did not (61% vs. 64%) as well as for students who indicated that they did not plan to enroll full-time as compared to those who did (59% vs. 63%). The somewhat counterintuitive result of students intending to live on campus having lower retention rates was largely due to those students being more likely to transfer to another institution (11% vs. 7%; adjusted OR = 1.6). The higher retention rates for students who planned to enroll full-time was largely attributed to those students being less likely to drop out of college (28% vs. 34%; adjusted OR = 0.8). For the

four-year sample, from a practical significance perspective, retention rates were somewhat comparable when examined by students' intentions of living on campus and enrolling full-time. For both samples, comparable retention and attrition rates were also observed by students' educational plans.

Retention rates increased as the number of students' college preferences met by their initial institution increased. The effect for this predictor was larger for the two-year sample than for the four-year sample (56% vs. 67% for the two-year sample and 76% vs. 80% for the four-year sample comparing met 0 vs. 3 preferences, respectively). Students who initially enrolled in an institution that matched fewer of their college preferences were more likely to drop out of college (adjusted OR = 1.1 to 1.6), as well as to transfer to another college in comparison to returning to their initial institution (adjusted OR = 1.1 to 1.7).¹⁸ Moreover, students who attended a college that was closer to home were more likely to return to their initial institution than those who attended a college farther away primarily because those who lived closer to home were less likely to transfer to another institution (adjusted OR = 0.6 to 0.9 for the four-year sample and adjusted OR = 0.3 to 0.7 for the two-year sample based on the transformed distance variable). This latter finding is further illustrated in Figure 2 for the four-year sample and Figure 3 for the two-year sample, where modeled retention and attrition rates are shown by the transformed distance from home variable on a continuous scale.

¹⁸ Multiple-predictor models were also estimated that included indicators for whether or not each individual college preference was met by the initial institution attended in place of using the overall variable of the number of preferences met. For the four-year sample, all three of the individual indicators were statistically significant at the 0.01 significance level and suggested that a match on the specific college preference was associated with lower dropout and transfer rates at the 0.01 significance level. For the two-year sample the indicators for state preference and institution type were statistically significant but the indicator for institution size was not. For both samples, a slightly larger difference in dropout rates was associated with the state preference indicator than with the other two indicators, while a slightly larger difference in transfer rates was associated with the institution type indicator.

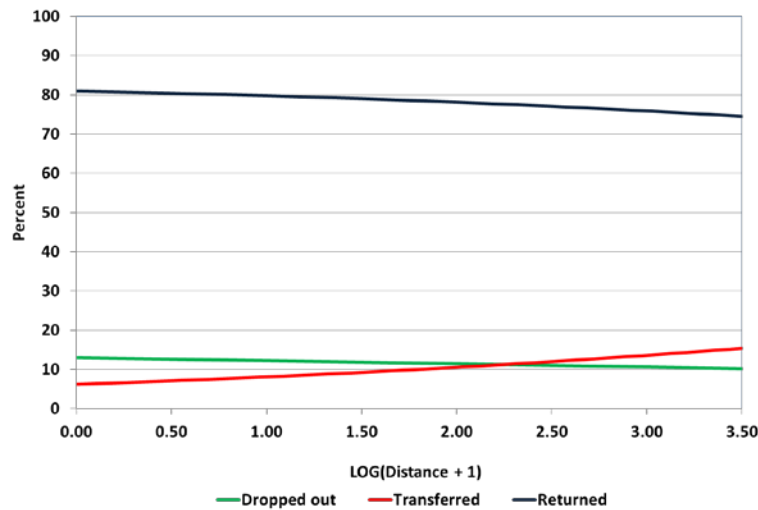


Figure 2. Modeled retention and attrition rates by transformed distance from home variable for the four-year sample, holding all other predictors constant at sample means¹⁹

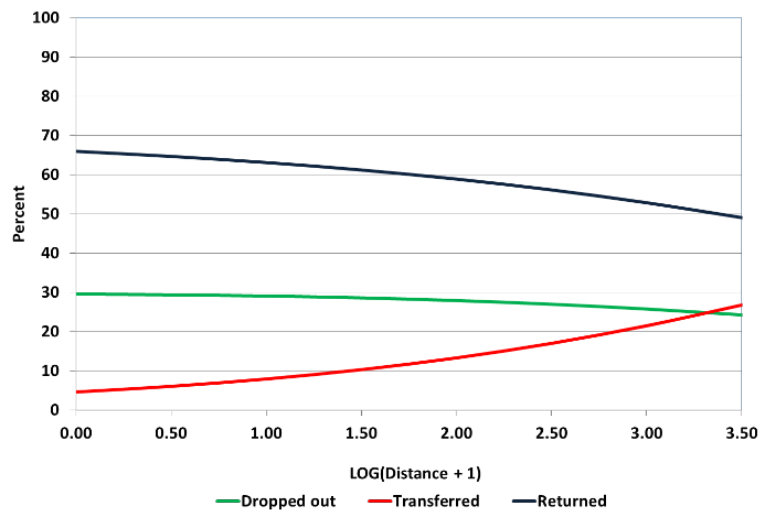


Figure 3. Modeled retention and attrition rates by transformed distance from home variable for the two-year sample, holding all other predictors constant at sample means¹⁹

Results by student demographic characteristics suggested that a higher percentage of females than males returned to their initial institution for both samples, because females were less likely than males to drop out of college (adjusted OR = 0.8 to 0.9; Tables A-1 and A-2).

¹⁹ Distance from home is 0 miles, 9 miles, 99 miles, and 999 miles when the transformed distance variable is 0, 1, 2, and 3, respectively.

Retention rates were slightly higher for students from more affluent neighborhoods than those from less affluent neighborhoods for both samples (79% vs. 76% for the four-year sample and 64% vs. 62% for the two-year sample). The higher retention rates were primarily attributed to students from less affluent neighborhoods being more likely to drop out in comparison to returning to their initial institution (adjusted OR = 1.1 to 1.3). For the two-year sample, students from less affluent neighborhoods were also slightly less likely to transfer to another institution in comparison to returning to their initial institution (adjusted OR = 0.8 to 0.9).

Compared to students whose parents earned a graduate degree, students whose parents had no college experience or some college experience were more likely to drop out of college compared to returning to their initial institution (adjusted OR = 1.3 to 1.4 for no college experience and adjusted OR = 1.2 for some college experience). This result was seen for both samples. Moreover, for the two-year sample, students whose parents had no or some college experience were slightly less likely than those whose parents earned a graduate-level degree to transfer to another institution in comparison to returning to their initial institution (adjusted OR = 0.7 to 0.8).

Results by race/ethnicity from the multiple-predictor models suggested that American Indian, Pacific Islanders, and multiracial students were less likely than White students to return to their initial institution because they generally had higher dropout rates (adjusted OR = 1.2 to 1.5 for dropped out vs. returned for both samples). In contrast, Asian and Hispanic students generally had higher retention rates than White students because they were less likely to drop out of college (adjusted OR = 0.8 to 0.9 for Hispanic students and adjusted OR = 0.6 to 0.8 for Asian students). For the two-year sample, African American students were more likely than White students to drop out of college in comparison to returning to their initial institution (adjusted OR

= 1.3). Retention and attrition rates were more comparable between African American students and White students for the four-year sample. It is important to keep in mind that these demographic comparisons were taken from the multiple-predictor models that accounted for students' achievement levels.²⁰

Results for Secondary Analyses

Rates by college major. Table A-4 provides the distributions for college major clusters by sample. For the four-year sample, the two major clusters with the highest percentage of students were Liberal Arts/General Studies (21%) and Business, Management, Marketing, & Related Services (12%). The two most predominant major clusters for the two-year sample were Liberal Arts/General Studies (55%) and Health Professionals (11%).

Retention and attrition rates varied across college major clusters (Figure 4 and Table A-5). For the four-year sample, the typical first-to-second year retention rate across college major clusters was 78% as compared to 76% overall irrespective of major; the retention rate ranged from 73% to 87% across major clusters.²¹ For the four-year sample, there was greater variability in transfer rates compared to dropout rates. For the two-year sample, the typical retention rate was 65% across college major clusters as compared to 60% overall irrespective of major; the retention rate ranged from 56% to 78% across major clusters.²² The dropout rate ranged from 18% to 32% across major clusters, while the transfer rate ranged from 3% to 15%.

²⁰ The reported demographic group differences in retention rates were reduced when students' academic preparation and achievement levels were taken into account as compared to the results from the single-predictor models.

²¹ Thirty-four out of the 36 major clusters were included in the four-year sample analyses. The two major clusters with fewer than 100 students that were deleted included: Science Technologies (41) and Construction Trades (46).

²² Twenty-six out of the 36 major clusters were included in the two-year sample analyses. The following ten major clusters had fewer than 100 students and were deleted from the analyses: Architecture & Related Services (04), Area, Ethnic, Cultural, Gender, & Group Studies (05), Foreign Languages, Literatures, & Linguistics (16), English Language & Literature (23), Mathematics & Statistics (27), Philosophy & Religious Studies (38), Theology & Religious Vocations (39), Science Technologies (41), Transportation & Materials Moving (49), and History (54).

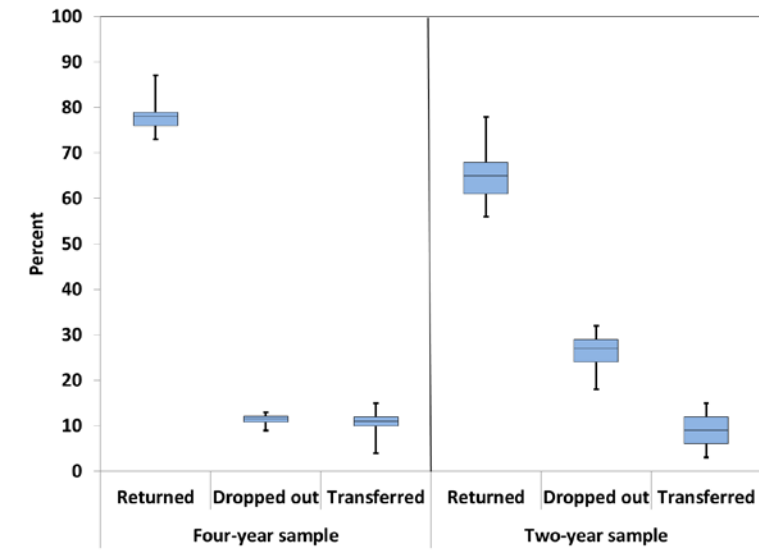


Figure 4. Modeled retention and attrition rates by college major clusters and sample, holding all other student-level predictors constant at sample means²³

Rates by institution characteristics. For the four-year sample, institution-level characteristics were also found to be significantly related to first-to-second year retention when these characteristics were added to the multiple-predictor models. The institution-level characteristics included: the control of the institution, HBCU status, institution size, and admissions selectivity. Institution size (the only institutional characteristic evaluated for the two-year sample) was not significantly related to retention at the 0.01 level for the two-year sample.

Table 9 provides the modeled retention and attrition rates and adjusted ORs for the dropped out vs. returned and the transferred vs. returned comparisons by institutional characteristics for the four-year sample, after statistically controlling for all other predictors in the model at their sample means. Retention rates were slightly higher at private institutions than at public institutions and at HBCU institutions than at non-HBCU institutions.

²³ The five parts of the boxplot include the minimum, the 25th percentile, median (50th percentile), 75th percentile, and maximum.

Table 9. *Results for First-to-Second Year Retention by Institutional Characteristic for the Four-Year Sample*

Institution characteristics	Modeled rates			Dropped out vs. returned			Transferred vs. returned		
	Dropped out	Transferred	Returned	OR	99% CI		OR	99% CI	
Control									
Private	11	10	79	0.83	0.74	0.94	0.99	0.88	1.10
<i>Public</i>	13	10	77						
HBCU									
Yes	11	8	80	0.91	0.75	1.09	0.77	0.64	0.92
<i>No</i>	12	10	78						
Size									
Under 1,000	14	12	74	1.30	1.01	1.67	1.70	1.33	2.16
1,000 – 4,999	12	11	77	1.08	0.91	1.28	1.44	1.23	1.70
5,000 – 9,999	12	10	78	1.11	0.95	1.30	1.33	1.14	1.54
10,000 – 19,999	12	9	79	1.06	0.91	1.24	1.10	0.95	1.27
<i>20,000 and above</i>	11	8	81						
Admissions selectivity									
Highly selective	10	5	85	0.56	0.45	0.70	0.41	0.33	0.51
Selective	10	10	80	0.59	0.49	0.71	0.85	0.71	1.01
Traditional	12	12	75	0.76	0.64	0.91	1.12	0.95	1.32
Liberal	13	12	75	0.81	0.66	1.00	1.08	0.88	1.32
<i>Open</i>	16	10	74						

Note. Italics indicate referent group. Adjustment was made for all student characteristics included in Table 7 and Table A-1. OR = odds ratio; CI = confidence interval. Characteristics for the postsecondary institutions were obtained from IPEDS, except for admissions selectivity. Admission selectivity was self-reported by institutions on the ACT IDQ as defined by the typical high school class ranks of their accepted freshmen: the majority of freshmen at highly selective schools are in the top 10%, selective in the top 25%, traditional in the top 50%, liberal in the top 75% of their high school class (ACT, 2015b). Institutions with open admissions policies accept all high school graduates to the limit of capacity.

As the size of the institution increased, the typical retention rate increased from 74% for institutions with under 1,000 students to 81% for institutions with 20,000 and more students. Compared to those enrolled at the largest institutions, students at smaller institutions were not only more likely to drop out of college (adjusted OR = 1.1 to 1.3), but they were also more likely to transfer to another institution (adjusted OR = 1.1 to 1.7). Additionally, as an institution's admissions policy became more selective, the typical retention rate increased from 74% for open

institutions to 85% for highly selective institutions. Students at highly selective institutions, selective institutions, and traditional institutions were less likely than those at open four-year institutions to drop out of college (adjusted OR = 0.6 to 0.8). Moreover, students at highly selective institutions were less likely than those at open four-year institutions to transfer to another college (adjusted OR = 0.4).

Discussion

Given the increased pressure that postsecondary institutions are under to improve their retention and degree completion rates, institutions are interested in determining student information that can help supplement their early alert warning systems and identify students who are at risk of leaving their institution (e.g., Tampke, 2013). To assist in this area, this study sought to identify some incoming student information from the ACT record that institutions might find helpful in determining early which students are at-risk of leaving their institution by either dropping out of college or transferring to another institution. This study found that student attrition between the first and second year at both two- and four-year institutions was not only related to academic readiness and demographic characteristics, but was also associated with students' college intentions, number of college preferences met by the initial institution attended, and distance from home. The benefits of using pre-enrollment information instead of waiting until midterm grades from the first term are available allows institutions to identify early students who may be more likely to leave the institution so that these students can be advised at college entry about the various academic and student support services that are available.

Findings related to the academic readiness measures and demographic characteristics were consistent with those reported elsewhere in the literature, including the finding that accounting for academic preparation helps to reduce the gaps among demographic groups (Kopp

& Shaw, 2016; Mattern & Patterson, 2009; Radunzel & Noble, 2013). Specifically, the current study found that students beginning at both two- and four-year institutions who were better prepared academically (as measured by standardized test scores and the coursework taken and grades earned in high school) were less likely to drop out of college than were those who entered underprepared. In terms of transferring to another institution, better prepared students beginning at four-year institutions were less likely to transfer, but those beginning at two-year institutions were slightly more likely to do so. This latter finding makes sense as it has been suggested that about one-third of students beginning at a two-year institution go on to transfer to a four-year institution within six years of initially enrolling in college (Jenkins & Fink, 2016). Moreover, it is the higher-achieving students that have been found to be more likely to transfer from a two-year to a four-year institution (e.g., Porchea, Allen, Robbins, & Phelps, 2010).

Although gaps in retention rates by demographic groups were reduced after statistically controlling for academic readiness measures, some of the gaps persisted especially when evaluated by parental education level and median household income. Students whose parents had no college experience (i.e., first-generation students) and those who came from less affluent neighborhoods were more likely than their peers to drop out of college. Among those beginning at two-year institutions, they were also less likely to transfer to another institution. In other studies, students beginning at a two-year institution from these specific demographic groups have been found to be less likely to go on to transfer to a four-year institution (e.g., Horn & Skomsvold, 2011). Based on these findings, institutions may want to have special programs in place that equip first-generation and low-income students with the resources and tools they need to succeed in college. Given that first-generation and low-income students often lack the guidance and support at home that can help contribute to their success in college (Saenz et al.,

2007; Westbrook & Scott, 2012), some researchers have suggested that institutions should engage faculty and peers in mentoring students from these demographic groups to provide them with academic and social support (e.g., Institute for Higher Education, 2012; Engle, Bermeo, & O'Brien, 2006; Engle & Tinto, 2008).

In terms of students' college intentions, the number of hours a student plans to work while in college had the largest effect on student attrition among the three college intentions examined in this study. This finding was seen at both two- and four-year institutions where student dropout rates increased as the number of hours planned to work increased. Assuming that what students say they are going to do is related to what they actually end up doing (Ajzen, 1991), this finding is in general agreement with that reported in other studies that suggests that working many hours (e.g., more than 20 hours) is negatively related to student retention (Astin, 1984) and to the academic performance of first-year students (Pike, Kuh, & Massa-McKinley, 2008). Conversely, research from these same studies has also suggested that working on-campus for a moderate number of hours has been found to be positively related to student retention and academic performance as these students tend to become more socially integrated at the institution.

For the other two college intentions examined in the current study (i.e., of enrolling full-time and living on campus), their effects on student attrition were larger at two-year institutions than at four-year institutions. At both types of institutions, students with intentions of enrolling full-time were found to be slightly less likely to drop out of college than those with intentions of enrolling part-time. This finding is consistent with those reported in a recent study by the NSC Research Center (2016) that suggested that part-time students have lower retention rates than full-time students. In comparison to the NSC study that is based on actual full-/part-time

enrollment status, the current study found substantially smaller differences in retention rates between students with intentions of enrolling full- and part-time. Differences in retention rates between students living on campus and off campus when based on student intentions in this study for the four-year sample were also smaller than those based on those reported in other studies that are based on actual campus residency status. For example, a national longitudinal study by Schudde (2011) found a three percentage point difference in first-to-second year retention rates between on-campus and off-campus residents. In contrast, students in the current study beginning at two-year institutions with intentions of living on campus had lower first-to-second year retention rates because they were slightly more likely to transfer to another institution. Generally, on-campus housing is available at most four-year institutions, while such options are offered at only about one-fourth of two-year institutions (American Association of Community Colleges, 2013). Among the current study's sample of two-year students transferring at year two, the percentage of students transferring to a four-year institution (instead of to another two-year institution) was slightly higher for students with intentions of living on campus than for those without these intentions (64% vs. 59%).

Another finding in the current study was that as the number of college preferences met decreased, students' chances of dropping out as well as transferring to another institution both increased. These relationships were more pronounced at two-year institutions than at four-year institutions. The negative relationships are in line with Tinto's arguments (1975, 1993) that a mismatch between the institutional environment and students' interests, needs, and preferences can play a role in a student's decision to leave an institution. A recent study (Bowman & Denson, 2014) using a Student-Institution Fit instrument found that greater student-institution fit was directly related to increased college satisfaction and indirectly related to greater intentions of

persisting. Several dimensions of fit were examined by their instrument that was administered to college students; these dimensions covered the following aspects of the campus environment: religious, athletic, academic, socioeconomic, political, physical, and social. In the current study, students' college preferences provided at the time students registered to take the ACT were compared to the characteristics of the initial institution attended to derive the number of preferences met. While this is a limited measure of student-institution fit, the results of this study and those of another ACT study (2014) suggest that examining incoming students' college preferences may help institutions to identify early those who may be more likely to leave their institution.²⁴ Interestingly, among transfer students, the median number of college preferences met by the second institution was lower than the number met by the initial institution for the four-year sample (1 vs. 2 matches, respectively) but was higher for the two-year sample (2 vs. 1 matches).²⁵

Distance from home was another variable that was identified as being related to student attrition. Specifically, lower retention rates were seen for students who attended a college that was farther from home primarily because these students tended to be more likely to transfer to another institution. This finding is consistent with that reported by another study (Mattern et al., 2013) that focused on students beginning at four-year institutions. In this earlier study, the researchers also found that transfer students tended to relocate closer to home. In subsequent

²⁴ An earlier ACT study (2014) examined the number of preferences met on type, location, and distance, and similar to this study found a negative relationship between the number of student preferences met and the percentage of students who transferred to another institution. In the current study we did not include distance as one of the match characteristics because approximately one-fourth of students in the analysis sample responded that they had no particular college in mind when they were asked about how far away they planned to live from the college that they expected to attend.

²⁵ For transfer students from the four-year sample, the number of college preferences met by the second institution compared to the number met by the initial institution attended varied: the number increased for 19% of students, the number decreased for 44% of students, and the number remained the same for 37% of students. The corresponding percentages for transfer students from the two-year sample were 44%, 16%, and 40%, respectively.

descriptive analyses, we also found this to be the case for the four-year sample where the median distance from home was 98 miles for the initial institution as compared to 25 miles for the second institution.²⁶ For the two-year sample, transfer students tended to relocate a little farther from home (median = 20 miles from initial institution compared to 58 miles from second institution).²⁷ Relocating to an institution farther from home was seen for two-year students transferring to a four-year institution but not for those transferring to another two-year institution. Findings from these studies suggest that if institutions are not already considering distance from home that they may want to explore the utility of it in identifying students who may be at risk of leaving their institution. Distance from home is a variable that can be easily calculated for all incoming students using standard software that computes the geodetic distance in miles between two zip code locations; a student's home zip code is available on the ACT record.

Future research might include examining student retention in relation to students' college intentions, preferences, and distance from home in combination with other noncognitive attributes that institutions may have available on their students at the beginning of the academic year, as well as evaluating these possible predictors in relation to other longer-term outcomes of retention and academic performance through degree completion. Given that retention rates were found to vary by college major, future research might also include exploring whether the effects of student characteristics on student retention and other college outcomes depend on college

²⁶ The median difference in the distance from home between the initial institution and the second institution was -41 miles for transfer students from the four-year sample.

²⁷ The median difference in the distance from home between the initial institution and the second institution was 18 miles for transfer students from the two-year sample.

major. This information may provide additional insights on ways institutions can make their student resources and supports more personalized to better meet students' needs.

In conclusion, the study findings illustrate how institutions can use incoming student information from the ACT record to help identify students who are at-risk of leaving their institution, allowing for the opportunity to intervene early with these students. Specifically, we focused on data elements thought to serve as possible proxies for barriers to social integration at the initial institution attended, such as students' college intentions on living on campus, enrolling full-time, and number of hours planned to work; number of college preferences met based on type, size, and state location; and distance from home. The ACT student record contains many data elements including ones that were not examined in the current study (such as the ACT Interest Inventory scores and college extracurricular plans) that can help institutions build and/or augment their multidimensional models of student success in order to better identify students who might benefit from additional academic and student support services upon entering college. Additionally, information from the ACT record could be incorporated into student-level dashboards to help faculty advisors learn more about their incoming students and equip them to better serve their advisees.

References

- Achieve. (2008). *The building blocks of success: Higher-level math for all students*. (Achieve Policy Brief). Washington, DC: Author.
- ACT. (2010). *Mind the gaps: How college readiness narrows achievement gaps in college success*. Iowa City, IA: Author.
- ACT. (2013). *Readiness matters: The impact of college readiness on college persistence and degree completion*. Iowa City, IA: Author.
- ACT. (2014). *A profile of 2012 ACT-tested high school graduates college choice report: Part 3: Persistence and transfer*. Iowa City, IA: Author.
- ACT. (2015a). *Expanding opportunities: A college choice report for the graduating class of 2014; Part 2: Enrollment patterns*. Iowa City, IA: Author.
- ACT. (2015b). *National collegiate retention and persistence-to-degree rates*. Iowa City, IA: Author.
- Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2), 179–211.
- Allen, J., Robbins, S. B., Casillas, A., & Oh, I. S. (2008). Third-year college retention and transfer: Effects of academic performance, motivation, and social connectedness. *Research in Higher Education*, 49(7), 647–664.
- American Association of Community Colleges. (2013). *Data points: On-campus housing*. Washington, DC: Author. Retrieved from http://www.aacc.nche.edu/Publications/datapoints/Documents/CampusHouse_8.28.13_final.pdf.
- Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel*, 25, 297–308.
- Bowen, W. G., Chingos, M. M., & McPherson, M. S. (2009). *Crossing the finish line: Completing college at America's public universities*. Princeton, NJ: Princeton University Press.
- Bowman, N. A., & Denson, N. (2014). A missing piece of the departure puzzle: Student–institution fit and intent to persist. *Research in Higher Education*, 55(2), 123–142.

- Bradburn, E. M. (2002). *Short-term enrollment in postsecondary education: Student background and institutional differences in reasons for early departures, 1996-98*. (NCES 2003-153). Washington, DC: National Center for Education Statistics.
- Brooks, J. H., II, & DuBois, D. L. (1995). Individual and environmental predictors of adjustment during the first year of college. *Journal of College Student Development*, 36(4), 347–360.
- D’Amico, M. M., & Dika, S. L. (2013). Using data known at the time of admission to predict first-generation college student success. *Journal of College Student Retention: Research, Theory & Practice*, 15(2), 173-192.
- Engle, J., Bermeo, A., & O’Brien, C. (2006). *Straight from the source: What works for first-generation college students*. Washington, DC: The Pell Institute for the Study of Opportunity in Higher Education.
- Engle, J., & Tinto, V. (2008). *Moving beyond access: College success for low-income, first-generation students*. Washington, DC: The Pell Institute for the Study of Opportunity in Higher Education.
- Fisher, S., Murray, K., & Frazer, N. A. (1985). Homesickness, health and efficiency in first-year students. *Journal of Environmental Psychology*, 5, 181–195.
- Horn, L., & Skomsvold, P. (2011). *Community college student outcomes: 1994-2009. Web tables (NCES 2012-253)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Institute for Higher Education Policy. (2012). *Supporting first-generation college students through classroom-based practices*. Washington, DC: Author.
- Ishitani, T. T. (2016). Time-Varying Effects of Academic and Social Integration on Student Persistence for First and Second Years in College National Data Approach. *Journal of College Student Retention: Research, Theory & Practice*, DOI: 10.1177/1521025115622781.
- Jenkins, D., & Fink, J. (2016). *Tracking transfer: New measures of institutional and state effectiveness in helping community college students attain bachelor’s degrees*. New York, NY: Community College Research Center.
- Kena, G., Hussar W., McFarland J., de Brey C., Musu-Gillette, L., Wang, X., ... Dunlop Velez, E. (2016). *The Condition of Education 2016* (NCES 2016-144). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Kopp, J. P., & Shaw, E. J. (2016). How final is leaving college while in academic jeopardy? Examining the utility of differentiating college leavers by academic standing. *Journal of College Student Retention: Research, Theory, & Practice*, 18(1), 2–30.

- Kuh G. D., Kinzie J., Buckley J. A., Bridges B. K., & Hayek J. C. (2006). *What matters to student success: A review of the literature* (Report commissioned for the National Symposium on Postsecondary Student Success: Spearheading a Dialog on Student Success). Washington, DC: National Postsecondary Education Cooperative.
- Lotkowski, V. A., Robbins, S. B., & Noeth, R. J. (2004). *The Role of Academic and Non-Academic Factors in Improving College Retention*. (ACT Policy Report). Iowa City, IA: ACT.
- Mattern, K. D., & Patterson, B. F. (2009). *Is performance on the SAT related to college retention?* (College Board Research Report No. 2009-7). New York, NY: The College Board.
- Mattern, K. D., Wyatt, J. N., & Shaw, E. J. (2013). College distance from home: Implications for student transfer behavior. *Journal of the First-Year Experience & Students in Transition*, 25(1), 77–92.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). New York, NY: Academic Press.
- National Student Clearinghouse Research Center. (2016). *Snapshot report – persistence and retention*. Herndon, VA: Author. Retrieved from <https://nscresearchcenter.org/wp-content/uploads/SnapshotReport18-PersistenceRetention.pdf>.
- Pike, G. R., Kuh, G. D., & Massa-McKinley, R. C. (2008). First-year students' employment, engagement, and academic achievement: Untangling the relationship between work and grades. *NASPA Journal*, 45(4), 560–582.
- Porchea, S. F., Allen, J., Robbins, S., & Phelps, R. P. (2010). Predictors of long-term enrollment and degree outcomes for community college students: Integrating academic, psychosocial, socio-demographic, and situational factors. *The Journal of Higher Education*, 81(6), 750–778.
- Radunzel, J., & Noble, J. (2012). *Tracking 2003 ACT-tested high school graduates: college readiness, enrollment and long-term success* (ACT Research Report No. 2012-2). Iowa City, IA: ACT.
- Saenz, V. B., Hurtado, S., Barrera, D., Wolf, D., & Yeung, F. (2007). *First in my family: A profile of first-generation college students at four-year institutions since 1971*. Los Angeles: Higher Education Research Institute, UCLA.
- Sanchez, E., & Buddin, R. (2016). *How accurate are self-reported high school courses, course grades, and grade point average?* (ACT Research Report No. 2016-3). Iowa City, IA: ACT.

- Schudde, L. T. (2011). The causal effect of campus residency on college student retention. *Review of Higher Education, 34*(4), 581–610.
- Shaw, E. J., & Mattern, K. D. (2009). *Examining the accuracy of self-reported high school grade point average*. (College Board Research Report 2009-5). New York: College Board.
- Tampke, D. R. (2013). Developing, implementing, and assessing an early alert system. *Journal of College Student Retention: Research, Theory, and Practice, 14*(4), 523–532.
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*, 89–125.
- Tinto, V. (1993). *Leaving college: Rethinking the cause and cures of student attrition* (2nd ed.). Chicago: The University of Chicago Press.
- Tinto, V. (2012). *Completing college: Rethinking institutional action*. Chicago: The University of Chicago Press.
- Tognoli, J. (2003). Leaving home: Homesickness, place attachment, and transition among residential college students. *Journal of College Student Psychotherapy, 18*, 35–48.
- Westbrook, S. B., & Scott, J. A. (2012). The influence of parents on the persistence decisions of first-generation college students. *Focus on Colleges, Universities, and Schools, 6*(1), 1–9.
- Wohlgemuth, D., Whalen, D., Sullivan, J., Nading, C., Shelley, M., & Wang, Y. (2007). Financial, academic, environmental influences on the retention and graduation of students. *Journal of College Student Retention, 8*(4), 457–475.

Appendix A

Table A-1. *Multivariate Results for First-to-Second Year Retention by Student Demographic Characteristics for the Four-Year Sample*

Student characteristics	Modeled rates			Dropped out vs. returned			Transferred vs. returned		
	Dropped out	Transferred	Returned	OR	99% CI		OR	99% CI	
Gender									
Female	10	10	80	0.76	0.74	0.78	0.94	0.91	0.96
<i>Male</i>	13	10	77						
Race/ethnicity									
African American	12	9	79	1.05	1.01	1.09	0.84	0.80	0.88
American Indian	16	11	74	1.47	1.30	1.66	1.05	0.90	1.23
Hispanic	11	10	80	0.92	0.88	0.96	0.89	0.85	0.93
Asian	9	8	82	0.76	0.71	0.81	0.73	0.68	0.79
Pacific Islander	14	12	74	1.28	1.01	1.61	1.17	0.89	1.52
Multiracial	14	11	76	1.21	1.15	1.28	1.04	1.02	1.09
<i>White</i>	12	11	78						
Median household income*									
< \$43,316	13	10	76	1.32	1.29	1.37	1.04	1.00	1.08
\$43,316 to \$61,580	12	10	78	1.12	1.09	1.16	1.02	0.99	1.05
> \$61,580	11	10	79						
Highest parental education level									
No college	14	10	76	1.39	1.34	1.44	1.05	1.00	1.10
Some college	13	11	77	1.24	1.20	1.28	1.14	1.10	1.18
Bachelor's degree	10	10	79	0.99	0.96	1.02	1.04	1.00	1.07
<i>Graduate degree</i>	11	10	80						

Note. Italics indicate referent group. Adjustment was made for all student characteristics in Table 7, as well as for all the demographic characteristics included in this table. OR = odds ratio; CI = confidence interval.

* Median household income is based students' residential zip code.

Table A-2. *Multivariate Results for First-to-Second Year Retention by Student Demographic Characteristics for the Two-Year Sample*

Student characteristics	Modeled rates			Dropped out vs. returned			Transferred vs. returned		
	Dropped out	Transferred	Returned	OR	99% CI		OR	99% CI	
Gender									
Female	27	9	64	0.87	0.84	0.91	1.01	0.95	1.07
<i>Male</i>	30	8	61						
Race/ethnicity									
African American	34	9	57	1.32	1.24	1.40	1.18	1.07	1.29
American Indian	36	9	56	1.43	1.18	1.75	1.12	0.81	1.54
Hispanic	25	8	67	0.84	0.79	0.90	0.86	0.77	0.96
Asian	19	8	73	0.57	0.50	0.67	0.78	0.63	0.97
Pacific Islander	32	6	61	1.18	0.78	1.76	0.77	0.37	1.60
Multiracial	34	7	59	1.27	1.16	1.40	0.92	0.79	1.08
<i>White</i>	28	9	63						
Median household income*									
< \$43,316	31	8	62	1.20	1.13	1.28	0.81	0.74	0.88
\$43,316 to \$61,580	28	9	63	1.08	1.03	1.15	0.87	0.81	0.94
> \$61,580	26	10	64						
Highest parental education level									
No college	33	7	61	1.34	1.25	1.44	0.67	0.60	0.74
Some college	29	8	62	1.18	1.10	1.26	0.79	0.73	0.86
Bachelor's degree	26	10	64	1.00	0.94	1.07	0.93	0.85	1.01
<i>Graduate degree</i>	25	11	64						

Note. Italics indicate referent group. Adjustment was made for all student characteristics in Table 8, as well as for all the demographic characteristics included in this table. OR = odds ratio; CI = confidence interval.

* Median household income is based students' residential zip code.

Table A-3. *Random Intercept Variance Estimates for Retention Outcome by Study Sample*

Model	Dropped out vs. Returned				Transferred vs. Returned			
	Variance estimate	Standard error	Range across institutions		Variance estimate	Standard error	Range across institutions	
			Min	Max			Min	Max
Four-year sample								
Null	0.461	0.024	-3.362	2.955	0.437	0.024	-4.152	2.802
Multivariate	0.226	0.013	-2.852	2.251	0.266	0.016	-3.705	2.508
Two-year sample								
Null	0.111	0.012	-1.428	0.735	0.230	0.024	-2.770	1.085
Multivariate	0.109	0.012	-1.630	0.600	0.192	0.021	-3.008	1.407

Note. The multivariate model includes the student characteristics presented in Tables 7 and A-1 for the four-year sample and in Tables 8 and A-2 for the two-year sample. The student characteristics were grand mean centered in the multivariate models.

Table A-4. *Description of College Major Clusters by Study Samples*

College major cluster	Four-year sample (N=445,115)		Two-year sample (N=84,674)	
	<i>n</i>	Pct.	<i>n</i>	Pct.
Agriculture, Agriculture Operations, & Related Sciences (01)	6,833	2	906	1
Natural Resources & Conservation (03)	2,623	1	111	<1
Architecture & Related Services (04)	1,791	<1	54	<1
Area, Ethnic, Cultural, Gender, & Group Studies (05)	335	<1	4	<1
Communication & Journalism (09)	13,358	3	263	<1
Communications Technologies/Technicians (10)	536	<1	304	<1
Computer & Information Sciences (11)	9,722	2	1,396	2
Personal & Culinary Services (12)	457	<1	863	1
Education (13)	25,033	6	2,345	3
Engineering (14)	37,627	8	1,436	2
Engineering Technologies (15)	3,513	1	2,023	2
Foreign Languages, Literatures, & Linguistics (16)	1,459	<1	65	<1
Family & Consumer Sciences (19)	3,329	1	745	1
Legal Professions & Studies (22)	606	<1	139	<1
English Language & Literature (23)	4,796	1	85	<1
Liberal Arts & General Studies (24)	92,110	21	46,668	55
Biological & Biomedical Sciences (26)	39,611	9	633	1
Mathematics & Statistics (27)	3,422	1	92	<1
Multi/Interdisciplinary Studies (30)	8,165	2	2,985	4
Parks, Recreation, Leisure, & Fitness Studies (31)	12,515	3	234	<1
Philosophy & Religious Studies (38)	653	<1	4	<1
Theology & Religious Vocations (39)	1,158	<1	7	<1
Physical Sciences (40)	9,538	2	154	<1
Science Technologies (41)	52	<1	96	<1
Psychology (42)	16,493	4	491	1
Homeland Security & Law Enforcement (43)	9,240	2	2,623	3
Public Administration & Social Services (44)	3,020	1	325	<1
Social Sciences (45)	12,497	3	287	<1
Construction Trades (46)	79	<1	501	1
Mechanic & Repair Technologies (47)	452	<1	1,640	2
Precision Production (48)	152	<1	813	1
Transportation & Materials Moving (49)	1,073	<1	51	<1
Visual & Performing Arts (50)	16,493	4	1,564	2
Health Professionals (51)	49,701	11	9,437	11
Business, Management, Marketing, & Related Services (52)	54,136	12	5,276	6
History (54)	2,537	1	54	<1

Note. College major clusters based on two-digit CIP codes. College major available for 84% of the four-year sample and 80% of the two-year sample.

Table A-5. *Description of College Major Clusters by Study Samples*

College major cluster	Four-year sample (N=445,115)			Two-year sample (N=84,674)		
	Dropped out	Trans- ferred	Returned	Dropped out	Trans- ferred	Returned
01	11	10	79	18	7	75
03	12	12	76	27	6	67
04	11	10	79	--	--	--
05	13	12	76	--	--	--
09	10	11	80	22	12	66
10	13	9	78	28	5	67
11	12	11	77	26	6	68
12	13	8	78	30	3	66
13	10	10	79	27	9	64
14	12	11	78	25	12	63
15	11	11	77	25	6	69
16	11	11	78	--	--	--
19	10	11	79	28	8	64
22	13	15	73	32	12	56
23	11	11	77	--	--	--
24	12	14	74	27	11	62
26	11	14	75	22	12	66
27	11	12	77	--	--	--
30	12	14	74	25	10	65
31	10	12	78	26	13	61
38	12	11	76	--	--	--
39	11	7	81	--	--	--
40	12	13	75	24	15	61
41	--	--	--	--	--	--
42	12	12	76	29	12	59
43	12	12	76	31	9	60
44	12	12	76	26	6	68
45	11	11	78	30	12	58
46	--	--	--	21	3	76
47	9	6	85	21	4	75
48	9	4	87	19	4	78
49	10	9	81	--	--	--
50	12	9	79	29	6	65
51	11	14	76	29	9	63
52	10	11	78	30	11	60
54	12	10	78	--	--	--

Note. College major clusters based on two-digit CIP codes. College major available for 84% of the four-year sample and 80% of the two-year sample. -- indicates that there were fewer than 100 students in the major.